Research Note 84-1

TASK-ELEMENT AND INDIVIDUAL DIFFERENCES IN PROCEDURAL LEARNING AND RETENTION: A MODEL-BASED ANALYSIS

Paul J. Sticha
Decisions and Designs, Inc.

C. Mazie Knerr
Human Resources Research Organization

Stephen L. Goldberg, Contracting Officer's Representative

Submitted by

Donald F. Haggard, Jr., Chief
ARI FIELD UNIT AT FORT KNOX, KENTUCKY

and

Harold F. O'Neil, Jr., Director
TRAINING RESEARCH LABORATORY

U. S. Army
Research Institute for the Behavioral and Social Sciences

January 1984

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Research Note 84-1	2. GOVT ACCESSION NO. A136789	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) TASK-ELEMENT AND INDIVIDUAL DIFFERENCES IN PROCEDURAL LEARNING AND RETENTION: A MODEL-BASED ANALYSIS		5. TYPE OF REPORT & PERIOD COVERED Final technical report Oct. 1981 to Feb. 1983
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Paul J. Sticha C. Mazie Knerr		8. CONTRACT OR GRANT NUMBER(s) MDA903-81-C-0517
9. PERFORMING ORGANIZATION NAME AND ADDRESS Decisions and Designs, Inc. Suite 600, 8400 Westpark Drive, P.O. Box 907 McLean, VA 22101		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2Q263743A794
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Institute Fort Knox Field Unit Fort Knox, KY 40121		12. REPORT DATE January 1984
		13. NUMBER OF PAGES 45
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Accession For	
NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A1	

DTIC COPY INSPECTED 2

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Mathematical models
Military training
Army training
Retention (Psychology)
Learning
Skills
SAINT (Systems Analysis of Integrated Networks of Tasks)
Armor tasks

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The rate at which performance improves during training, and the extent to which information is retained during intervals without practice, is a concern of those who plan and manage military training. This report was prepared for the U. S. Army Research Institute, Ft. Knox, Kentucky, and illustrates the application of mathematical models to investigate issues → over

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

regarding acquisition and retention of complex, military skills. The purpose of this report is to test a model of learning and retention of Armor procedures. Specifically, the ability of the model to account for task-element and individual differences identified in earlier research was examined.

The findings of this report provide some empirical support for a model of procedural skill learning and retention which could be used to assist the training manager in determining training requirements for various tasks. However, the analysis of learning and retention issues is largely exploratory, and future research is necessary to confirm the findings of this study. This report concludes with a discussion of possible directions future research could take.

Summary

Requirement

The purpose of this report is to test a model of learning and retention of Armor procedures. Specifically, the ability of the model to account for task-element and individual differences identified in earlier research was examined. In addition, this report illustrates how analytical models may be used to investigate issues in skill acquisition and retention.

Procedure

Soldiers from Armor One Station Unit Training (OSUT) were trained on two of eight procedural tasks from the OSUT Program of Instruction. The soldiers received five training trials on each task shortly after formal training for the task. A retention test was given one month later, at the time of the gate test for the task. Mathematical models of learning and retention were fit to the data. The models predicted differences in performance between task elements from ratings of five characteristics of the task elements, and individual differences from two composites of the Armed Services Vocational Aptitude Battery.

Findings

The mathematical models which accounted for task-element and individual differences provided a significantly better fit to the data than models which ignored these differences. Consideration of task-element differences produced a greater increase in the goodness-of-fit of the models than consideration of individual differences. The weights of the task-element characteristics and aptitude scores in predicting learning and retention parameters were not consistent across tasks, although there were some general trends in both analyses.

Use of Findings

The findings illustrate how mathematical models can be used to address issues related to acquisition and retention of skills. They also provide empirical support for a model of procedural skill learning and retention which could be used to assist the training manager in determining training requirements for various tasks. However, other issues, such as the theoretical prediction of parameter values, must be addressed before the model can be applied for this purpose.

Contents

Contents (Continued)

Tables

Figures

Task-Element and Individual Differences in
Procedural Learning and Retention:
A Model-Based Analysis

Introduction

It borders on a tautology that some things are easier to learn than others, and that on any particular task, some people learn faster, while others learn more slowly. Even within a single task, some of the individual steps are easier to learn or are retained better, while others are more difficult to master or are forgotten sooner. The rate at which performance improves during training, and the extent to which information is retained during intervals without practice, is a concern of those who plan and manage military training. For example, certain tasks or task elements which are difficult to learn require more training to achieve acceptable levels of performance; tasks which are forgotten quickly or are performed infrequently in the normal activities of a soldier require periodic retraining to ensure readiness.

Task-element and individual differences in acquisition and retention of military skills have been identified by a number of empirical studies. The results of this research have been summarized in several reviews (Wheaton, Rose, Fingerman, Korotkin, Holding, & Mirabella, 1976; Annett, 1977; Knerr, Berger, & Popelka, 1980). Much of the earlier research was performed in laboratory settings and used simple psychomotor or verbal learning tasks. The U. S. Army Research Institute (ARI) has examined these factors in military settings with Army technical tasks, and thus provides results pertinent to the present research; they are reviewed by Rose, McLaughlin, Felker, and Hagman (in press). Most recently, research for ARI has identified task-element and individual differences in a study which used mathematical models to investigate procedural learning and retention (Sticha, Edwards, & Patterson, 1982). The research described in this report follows on the results of Sticha et al., and attempts to characterize task-element and individual differences in terms of more basic variables.

The research described in this report also represents a methodological advancement in the investigation of military learning and retention issues. One purpose of this report is to illustrate some of the details of this new approach based on mathematical learning and retention models. The analysis is largely exploratory; thus, it will be necessary to confirm the findings of this study with future research. This report concludes with a discussion of the implications of the results and some of the possible directions future research could take.

Rationale for Model-Based Investigation

When the effects of task and individual variables on learning and retention are investigated, it is important that learning and retention are measured in a way that can meaningfully be compared across tasks and experimental conditions. Typically, learning is measured by the improvement in performance over a fixed number of training trials, or by the number of training trials required to achieve criterion performance. Retention is similarly measured by the difference in performance before and after an interval without practice.

These traditional measures of learning and retention are confounded by a number of variables which are not of primary interest to the researcher. Among these variables are the initial level of learning, the strictness of the performance criterion, and the time interval over which data are collected. Rose et al. (in press) have illustrated the problems that occur with simple measures of retention, because the rate of forgetting decreases over time. Research samples tested early in the curve, during rapid decay, show large amounts of forgetting, while samples tested later do not show decay.

The criticisms applied to the analysis of retention apply to acquisition, as well. The improvement in performance due to training is not linear, and simple measures of learning produce results that depend on details of the experimental procedure. For example, if two groups differ in the initial amount of learning, the group with greater initial learning would be expected to learn at a lower rate, even if the same learning curve applied to both groups.

In order to make meaningful statements about acquisition or retention, it is necessary to consider the entire learning or forgetting curve, which cannot be captured by sampling only two points from it. Mathematical models of acquisition or retention are an attempt to characterize learning and forgetting processes (that is, describe the shape of acquisition and retention curves) by a small number of parameters. If a model is successful, then statements made by the model about behavior will not vary with changes in exogenous variables.

A Model of Procedural Learning and Performance

A model describing the learning and performance of procedural tasks was developed and applied to eight Armor procedures by Sticha et al. (1982). This model combines a network representation of task-element sequencing with models of the psychological processes involved in acquisition, retention, retrieval,

and choice. The models were chosen based on a review of the modeling literature (Sticha, 1982), which considered criteria such as flexibility, validity, generality, and pragmatic concerns in evaluating modeling approaches.

Represenation of task-element sequencing. A framework for representing performance of the procedural tasks is provided by the SAINT (System Analysis of Integrated Networks of Tasks) simulation system (Wortman, Duket, Seifert, Hann, and Chubb, 1978 a,b). SAINT is a general system for discrete or continuous simulation of networks of tasks. Each step in a procedure is represented by a task in a SAINT model. The steps are linked in a network that represents the constraints on the orders in which the steps may be performed. Included in SAINT is the ability to reflect deterministic, probabilistic, and conditional branching between tasks, as well as more complex interactions in which tasks are modified by other tasks. The SAINT models are described in detail by Sticha et al. (1982).

Representation of psychological processes. Psychological models describing acquisition, retention, and retrieval, are represented in the overall model as subroutines within the SAINT system. The approach to learning and retention is based on the concept of the strength of an association. The strength of an association is assumed to be a normally distributed random variable. The probability of correctly retrieving the association is the probability that the strength of the association exceeds a threshold (Wickelgren, 1974b). Acquisition, according to this approach, is described by a function relating association strength to the amount of practice or number of training trials. The function used in the models follows the tradition of Hull (1943, 1952) in assuming that strength increases at a constant rate (that is, geometrically) to an asymptote.

The retention model describes the changes in strength of a memory trace that occur during intervals without practice. The model used follows the assumptions of strength-fragility theory (Wickelgren, 1974a), which postulates two processes that lead to loss of memory: a process that leads to very quick, exponential decay, and a process that leads to slower decay according to a power function. The long-term retention function represents a consolidation theory of memory dynamics. According to this theory, a new memory trace is fragile and decays at a rapid rate. As the memory trace gets older, the fragility decreases, and hence, the trace decays at a slower rate. Only the long-term component of strength-fragility theory was used in the models.

Validation of the models. The learning and retention models were validated by comparing their predictions to data gathered from two samples of soldiers: a sample of soldiers in One Station Unit Training (OSUT), and a sample of soldiers in an operational unit. The model offered a good account of the data from

the OSUT sample, and predicted the overall success rate, the average task-element success rate, and performance speed to a high level of accuracy. However, further analysis identified differences between task elements in the value of model parameters. Specifically, the ability of a model to predict the results could be improved significantly by estimating parameters separately for different task elements. In addition, the parameters estimated from one portion of the soldiers did not provide an optimal accounting for the data from the remaining soldiers, although the description was good. No attempts were made in these analyses to relate task-element or individual differences to more basic characteristics of the task-elements or individuals.

Analysis of data from the operational unit found no differences in performance as a function of the time since training in OSUT. It appears that the retention processes operating after initial training have reached their asymptote by the time soldiers were sampled for the experiment (at least 3 months) In addition, the results suggested that task elements differ in the extent to which they are practiced in the unit. These results suggest that future experiments investigating retention should be focused on repeated measures designs, naturalistic observation of performance, and documentation of training.

This report presents the results of a more detailed analysis of the data from the OSUT sample, and investigates the issues that were identified by Sticha et al. (1982) in the preliminary validation of the models. In particular, differences in performance between tasks and task elements will be related to task characteristics that have been shown to affect acquisition or retention, and individual differences in performance will be related to measures of aptitude.

Task-Element and Individual Differences in Learning and Retention

The analysis described in this report builds on a history of research in which task-element and individual differences in learning and retention have been documented in both military and academic settings. Research to identify task and individual variables which account for learning and retention differences has identified some variables, although our understanding of these differences is still incomplete.

Task variables. Schendel, Shields, and Katz (1978) succinctly state that "Procedural tasks and individual discrete motor responses are forgotten over retention intervals measured in terms of days, weeks, or months, whereas continuous movements typically show little or no forgetting over retention intervals measured in terms of months or years" (p. 5). The cognitive mechanism producing differences in retention of procedural and continuous tasks may be the extent of memorization, which is greater in procedural tasks. Most Army tasks, however, are

procedural, and thus, the global distinctions used to characterize tasks fail to distinguish the determinants of retention.

The differentiation of tasks into their components, skills, steps, or subtasks, leads to the detailed behavioral analysis of tasks to determine their stimuli, processes, and responses. These components, or subtasks, differ in their level of retention, as shown in existing research. Rose et al. (in press) summarize the types of tasks that have been examined in Army skill retention research, and note that descriptive analyses of the task and steps have been performed post hoc.

Dimensions of task steps and tasks that appear to reduce retention include the following:

1. Difficulty or high skill demand (Goldberg, Drillings, and Dressel, 1981; Osborn, Campbell, and Harris, 1979; McCluskey, Hiller, Bloom, and Whitmarch, 1978; Vineberg, 1975; Hagman, 1980 a & b),

2. Lack of cues from sequential steps, equipment, and so forth, often involving safety precautions (Goldberg et al., 1981; McCluskey et al., 1978; Osborn et al., 1979; Shields, Goldberg, and Dressel, 1979),

3. Unclear to the soldier or of questionable relevance to the task (Osborn et al., 1979; Shields et al., 1979),

4. First and last steps (Osborn et al., 1979),

5. Passive steps (Osborn et al., 1979),

6. Training and testing differences (Goldberg et al., 1981, Osborn et al., 1979), and

7. Interference from interpolated activities (Knerr, Harris, O'Brien, Sticha, and Goldberg, 1982).

Shields et al. (1979) and Knerr et al. (1982) also demonstrated that longer tasks (more steps) are learned more slowly and forgotten sooner than short tasks.

Individual differences. Aptitude differences influence skill acquisition and thus, indirectly influence retention. Army research demonstrates the favorable effects of general aptitude on skills in Air Defense and Field Artillery (Department of the Army, TRADOC Systems Analysis Activity [TRASANA], 1977; Field Artillery School, 1977). Rose et al. (in press) note, however, that Army research on the subject, as yet, is inconclusive.

Five projects conducted by the U.S. Army Research Institute (ARI) investigated the effects on skill retention of individual ability as measured by Army aptitude tests. Vineberg (1975) found a direct relationship between aptitude and performance on both initial and retention tests; however, the relationship did not hold for all tasks. Other ARI research discovered no significant relationships between aptitude and performance (Goldberg et al., 1981). The relationship may be mediated by training methods (Dressel, 1980; Holmgren et al., 1979; Sullivan, Casey, & Hebein, 1978).

Objectives

This research has three major objectives:

1. To provide a more detailed validation of the model of procedural learning and performance developed by Sticha et al. (1982).

2. To illustrate the application of mathematical models to the investigation of issues in the acquisition and retention of complex skills involved in military tasks.

3. To investigate characteristics which predict task-element and individual differences in learning and retention of Armor tasks.

The following sections present the approach that was used to meet these objectives, and presents and discusses the implications of the results.

Method

Task Selection

Procedural tasks were selected from those performed by the gunner, loader, or driver of an M60A1 tank. The following tasks were selected to represent a range of length, complexity, and extent of practice in the unit after initial training (values on these dimensions are reported by Knerr et al., 1982):

1. Load an M240 Machinegun (LOADMG)
2. Start the M60A1 Tank Engine (STARTANK)
3. Stop the M60A1 Tank Engine (STOPTANK)
4. Perform Gunner's Prepare-to-Fire Checks (GUNNERPF)
5. Perform Loader's Prepare-to-Fire Checks (LOADERPF)
6. Engage Targets using Precision Fire Techniques (PRECFIRE)
7. Communicate over Tactical FM Radio (RADIOMSG)
8. Communicate using Visual Signal Techniques (SIGNALS)

Behavioral Analysis

The tasks were analyzed to determine the task elements (steps), standards, and conditions of performance. The results of these analyses were used to develop test scenarios, score forms, and scorer training material.

Additional behavioral analyses of the task identified characteristics related to learning, performance, and retention. These characteristics were cast into questionnaire form, and rating booklets were compiled to gather ratings from project staff and noncommissioned officers who served as scorers in the data collection. Each task element was rated on the following fourteen characteristics:

1. Requires recall of knowledge
2. Requires rule learning and using
3. Requires guiding and steering, continuous movement
4. Has cues for performance
5. Has stimulus-response conflict
6. Has aversive consequences of failure
7. Has feedback
8. Step typically omitted in unit practice
9. Step performed differently in unit
10. Different step performed in unit practice
11. Step not performed in similar tasks
12. Difficult
13. Cricital to the overall performance of the task
14. Step performed in emergency or in combat

In the first seven of the task characteristics, the raters indicated the level of the factor for each task element by

making a mark on a line; the endpoints of the line were defined to be extreme levels of the characteristic. The marks were subsequently translated to a scale from 0 to 10. The scores of different raters were aggregated by taking the median. For the remaining task characteristics, raters stated whether the characteristic was present or absent. The score for a task element is the percentage of raters who judged that the characteristic was present for the task element.

Data Collection

The ability of the task characteristics and aptitude measures to predict procedural learning and retention was investigated using data from a sample of soldiers in Armor OSUT.

Subjects. Subjects were 471 soldiers from four OSUT companies at Ft. Knox, Kentucky in their fifth to tenth week of training.

Procedure. Each soldier performed two of the eight tasks for a total of six trials: five acquisition trials and a retention trial. For each task tested, the soldiers reported to the test site twice during a twelve-week data collection period with approximately four weeks between sessions. The first session coincided roughly with formal training of the task; the second session coincided roughly with the gate test for the task. Except for the fact that a task was performed five times in the first session, while it was performed only once in the second session, the procedure for the sessions was identical.

A session began by the scorer reading a set of instructions to inform the soldier of the task and any specific conditions to consider during performance (e.g., moving or stationary targets during precision fire engagements). After reading the instructions, the scorer did not intervene during the performance of the task unless the soldier made an error.

If the soldier committed an error on a step, the scorer gave him some assistance. If this degree of assistance was not sufficient to produce correct performance, the scorer gave stronger assistance, until correct performance on the step was obtained. The following three levels of assistance were used:

Level 1. Remind the soldier what the overall task is, and tell him the steps he has performed up to that point.

Level 2. Tell the soldier what the next step is.

Level 3. Show the soldier how to do the step.

After the soldier demonstrated the step correctly, he proceeded to the next step and continued until he had completed the task.

While the soldier performed the task, the scorer recorded data on correct performance of task steps, the order in which the soldier performed the steps, the type of error committed, the level of assistance given, and the elapsed time. Armed Service Vocational Aptitude Battery (ASVAB) scores and level of education were obtained from personnel records.

Results

Sticha et al. (1982) provide a preliminary analysis of the data in which differences in learning and retention between task elements and between individuals were identified as topics for further analysis. This analysis develops and tests models to investigate these issues.

Task-Element Differences

The basic learning and retention model has eight parameters, six of which are identifiable from the OSUT data. Three of these parameters are concerned with the acquisition component of the model: (1) initial strength, (2) strength asymptote, and (3) learning rate. Two parameters are present in the retention component of the model: (1) strength decay rate, and (2) fragility decay rate. Three parameters are present in the recall component of the model: (1) strength threshold for correct response, (2) strength threshold for first level of assistance, and (3) strength threshold for second level of assistance. Because the time between the fifth and sixth trials is constant within a task, there is no variation in the retention interval, and consequently, only six of the parameters can be estimated from the data. Specifically, either one of the thresholds, the initial strength, or the strength asymptote must be set arbitrarily, and either the strength decay rate or the fragility decay rate must be set arbitrarily. Thus, there are six free parameters to be estimated from the data.

The basic model that was tested by Sticha et al. (1982) pooled data from all task elements of a procedure to obtain parameter estimates for the procedure. Thus, the parameter values were assumed to be constant across task elements. Task-element differences were identified by comparing the fit of the basic learning and retention model to the fit of a model in which the task elements were divided into two groups, with separate parameters estimated for each group. Since the more complex model performed better than the basic model for six of the eight tasks, the hypothesis that learning and retention parameters were constant over task elements could be statistically rejected for those tasks. However, no attempt was made to relate differences in learning and retention parameters to task-element characteristics that could be independently assessed.

In this analysis, task-element differences were related to the task characteristics rated by members of the project staff and the noncommissioned officers who served as scorers. Each of the four acquisition and retention parameters (initial strength, learning rate, strength asymptote, and retention proportion) was assumed to be a linear function of five of the fourteen task characteristics: (1) extent of rule learning and

using; (2) aversivenss of consequences in covert performances; (3) degree of feedback; (4) extent of interference as measured by an index encompassing omission, differences in performance, performance of different steps, and performance of different steps in similar tasks; and (5) performance in emergency or combat. The interference index was ten times the sum of the four task characteristics relating to differences between the task as tested and as practiced. Other task characteristics were eliminated, because they did not have sufficient variance to produce reliable weights (Table 1). The interference characteristics were retained as a single index because of previous results using these data which indicated that retention was related to interference (Knerr et al., 1982).

The resulting model contains the following 26 free parameters: 3 thresholds; 3 constants for the linear equations for learning rate, strength asymptote, and decay rate (the constant for initial strength was set to 5.0); and 20 parameters representing the weights for the 5 independent variables predicting 4 dependent variables. The first 6 parameters correspond to the identifiable parameters in the basic acquisition and retention model. The remaining 20 parameters are the weights in the equations that predict parameter values from the ratings of the task characteristics. Previous experience with the models suggested that addition of parameters beyond this number would make optimization of parameter values too time-consuming.

The ability of the task characteristics to account for task-element differences was assessed by comparing the goodness-of-fit of the 26-parameter model described above with that of the basic 6-parameter model in which there are no task-element differences (for PRECFIRE and SIGNALS, the models have 27 and 7 parameters, respectively). When goodness-of-fit is measured by twice the negative log likelihood of the data given the model, the difference in the goodness-of-fit between the models has a chi-squared distribution with 20 degrees of freedom.

Parameters were estimated using an automated, iterative, unconstrained optimization routine. This routine starts with an initial set of parameters supplied by the user. A user-written subroutine then calculates the likelihood of the data given the current parameter values. The optimization routine then steps the parameters through a variety of values in order to find those values for which the likelihood of the data are maximized. As the parameter values get close to their optimal values, the size of the steps used to change parameter values is reduced until a criterion step size is obtained, and the optimal value of the parameters is returned. This solution must be examined to ensure that a global maximum was found, rather than a local maximum or boundary value.

The skill-rating models presented considerable difficulty to the optimization routine. To reduce these problems, the

Table 1
Means and Standard Deviations
of Skill Ratings

Rating	Mean (N=119)	Standard Deviation
Recall of Knowledge	3.70	1.82
Rule of Learning & Using	2.05	2.37
Guiding & Steering	0.82	1.29
Cues for Performance	0.65	1.75
Stimulus-Response Conflict	0.03	0.28
Aversive Consequences of Error	2.41	2.40
Feedback	5.04	3.70
Interference Index	4.22	3.43
Omission	0.15	0.19
Performs Different	0.08	0.13
Different Step	0.03	0.07
Not in Similar Tasks	0.17	0.11
Difficult	0.03	0.13
Critical	0.70	0.17
Performed in Combat	0.87	0.20

step size at which the optimization would stop was relaxed from the values used by Sticha et al. (1982). Although this change is probably not important, it may lead to slightly lower estimates of the degree of improvement for the skill-rating models. Table 2 shows the goodness-of-fit measure for the two models. For seven of the eight tasks, the improvement in prediction obtained by the model based on the skill ratings is large and highly significant. The difference represents an average 8.8% improvement in the fit of the model, with a range from 0.4% (SIGNALS) to 19.6% (LOADMG). The magnitude of task-element differences agrees with the results of Sticha et al. (1982).

Table 3 presents the weights of the skill components in the four linear equations predicting initial strength, learning rate, strength asymptote, and retention proportion for each of the eight tasks. To facilitate comparison of the weights across skill components, the weights were standardized by multiplying the raw weights by the standard deviation of the skill component over all tasks. Although use of the weights significantly improves model performance, it should be kept in mind that for some tasks (particularly LOADERPF and RADIOMSG) the number of task elements was close to the number of task characteristics, and hence, there may be extraneous sources of variation in the weights.

The weights show considerable variation across tasks. However, for some of the tasks, certain weights were quite high, often in a surprising direction, and deserve further discussion. For three tasks (LOADMG, STARTANK, and PRECFIRE), the interference index was positively related to initial strength, learning rate, and strength asymptote. This result is surprising for two reasons. First, interference is a variable which should primarily affect retention, rather than learning. Second, the effect of interference on learning, if any, would be expected to be in the other direction; that is, greater interference would be expected to lower the learning rate rather than raise it.

Aversive consequences and feedback do not have a consistent effect on the learning rate. Presence of feedback has a negative effect on four of the tasks and a positive effect on only one task. There was no variance in ratings of extent of feedback for SIGNALS. Aversive consequences, on the other hand, have a positive effect on four tasks and a negative effect on three tasks. There was no variance in task-element ratings for use of rules for three of the tasks (LOADERPF, RADIOMSG and SIGNALS). The value of the weights for this task characteristic, as well as for whether the task-element would be performed in combat, did not show any obvious trends.

The SAINT models of the eight tasks were run using the parameters of the task-characteristic model. One hundred simulated subjects were run for each task. The percentage of task-

Table 2
Goodness-of-Fit for Task-Element Difference Models

Task	Negative Log Likelihood: Single-Value Model	Negative Log Likelihood: Skill Rating Model	Chi-square for Improvement[a]
LOADMG	2021.54	1625.48	396.06*
STARTANK	2385.70	1983.01	402.69*
STOPTANK	2373.58	2107.86	265.72*
GUNNERPF	16613.20	15912.72	700.48*
LOADERPF	2706.74	2554.90	151.84*
PRECFIRE	9510.96	9043.96	467.00*
RADIOMSG	4291.52	3975.36	316.16*
SIGNALS	3990.58	3975.50	15.08

[a] df=20

* $p < .001$

Table 3
Standardized Skill-Component Weights

Task	Dependant Variable	Independent Variable Weights: Use of Rules	Aversive Consequences	Feedback	Interference	Performed in Combat
LOADMG	Initial Strength	-1.568	0.802	0.382	3.489	-0.295
	Learning Rate	0.0	0.0	0.062	0.429	-0.003
	Asymptote	-0.471	0.110	0.670	3.069	-0.347
	Retention	-0.017	0.038	0.032	0.134	0.053
STARTANK	Initial Strength	0.167	0.750	-0.778	0.409	0.202
	Learning Rate	-0.017	0.156	-0.302	0.004	0.043
	Asymptote	0.069	-0.689	-0.269	0.657	-0.035
	Retention	0.077	0.029	-0.053	-0.008	-0.027
STOPTANK	Initial Strength	-0.177	-0.491	0.356	-0.847	-1.077
	Learning Rate	-0.137	-0.096	0.0	0.101	-0.097
	Asymptote	2.544	4.190	-1.459	0.098	2.893
	Retention	-0.407	-0.282	0.216	-0.178	-0.442
GUNNERPF	Initial Strength	0.172	0.180	0.017	-0.760	0.092
	Learning Rate	0.001	-0.274	-0.105	0.356	0.553
	Asymptote	-0.023	1.174	0.164	-1.345	-1.576
	Retention	-0.005	-0.074	-0.006	0.055	0.065
LOADERPF	Initial Strength	0.0	-0.523	-0.098	-0.009	-0.419
	Learning Rate	0.0	0.631	0.0	0.0	0.353
	Asymptote	0.0	-1.652	-0.732	3.656	1.118
	Retention	0.0	0.005	0.097	-0.277	-0.195
PRECFIRE	Initial Strength	-0.081	-0.098	0.589	1.598	1.139
	Learning Rate	-0.137	0.040	-0.163	0.262	0.091
	Asymptote	1.384	-1.515	2.918	6.160	0.006
	Retention	-0.014	0.009	-0.011	-0.007	0.001
RADIOMSG	Initial Strength	0.0	-0.948	0.223	-0.090	0.168
	Learning Rate	0.0	0.055	-0.0002	0.0002	0.060
	Asymptote	0.0	-5.747	-1.465	0.813	0.027
	Retention	0.0	0.127	0.011	-0.021	-0.443
SIGNALS	Initial Strength	0.0	0.069	0.0	-0.038	0.090
	Learning Rate	0.0	-0.086	0.0	-0.056	-0.026
	Asymptote	0.0	-0.013	0.0	-0.200	0.080
	Retention	0.0	-0.009	0.0	0.005	0.009

elements performed correctly was compared to the data from the soldiers, as well as to the predictions of the basic six-parameter model (from Sticha et al., 1982). The performance for each task by trial and task element is plotted in Figures 1-8. The figures show the extent to which consideration of the five task characteristics improves the performance of the model.

The improvement brought about by the task-characteristic model is especially evident on the first trial, and in some cases, the retention trial. Even though the fit is impressive, there are some relatively large differences which are not predicted by the task-element model. For example, the final task element in LOADMG (Figure 1) exhibits very low performance which is not predicted by the task-characteristic model. Results of the PRECFIRE model, (Figure 6), illustrate the fact that the task characteristics do not capture all of the variance in performance. Task elements 8-11 all involve laying the cross-hair on a target with the proper lead. These task elements all received the same ratings on all task characteristics. Yet, there is considerable variance in performance, particularly on the first trial and the retention trial. Thus, additional factors, such as the soldiers familiarity with different kinds of ammunition, must be considered to account for task-element differences in learning and retention.

In summary, a model which predicts task-element differences as a function of five task characteristics provided a significant improvement over a model which assumed all task elements had the same values for the learning and retention parameters. However, the weights by which the task characteristics were combined to predict learning and retention were, in general, not consistent across tasks. In addition, some details of task-element performance were not predicted by consideration of the task characteristics.

Task Differences

The analysis of task-element differences suggests that, although task characteristics may account for the differences between task elements within a single task, the relationship is probably not consistent over tasks. This result is not entirely surprising for three reasons. First, on some tasks there are almost as many task characteristics as task elements. In this situation, the relationship between the model parameters and the task characteristics would include some variation that would be otherwise be counted as error, if there were a larger sample of task elements. These task-specific characteristics would lead to a relationship that varied over tasks. Second, it cannot be assumed that the task characteristics are measured at anything greater than an ordinal scale. The measuring device may be differentially sensitive to changes at different parts of the range for some task characteristics,

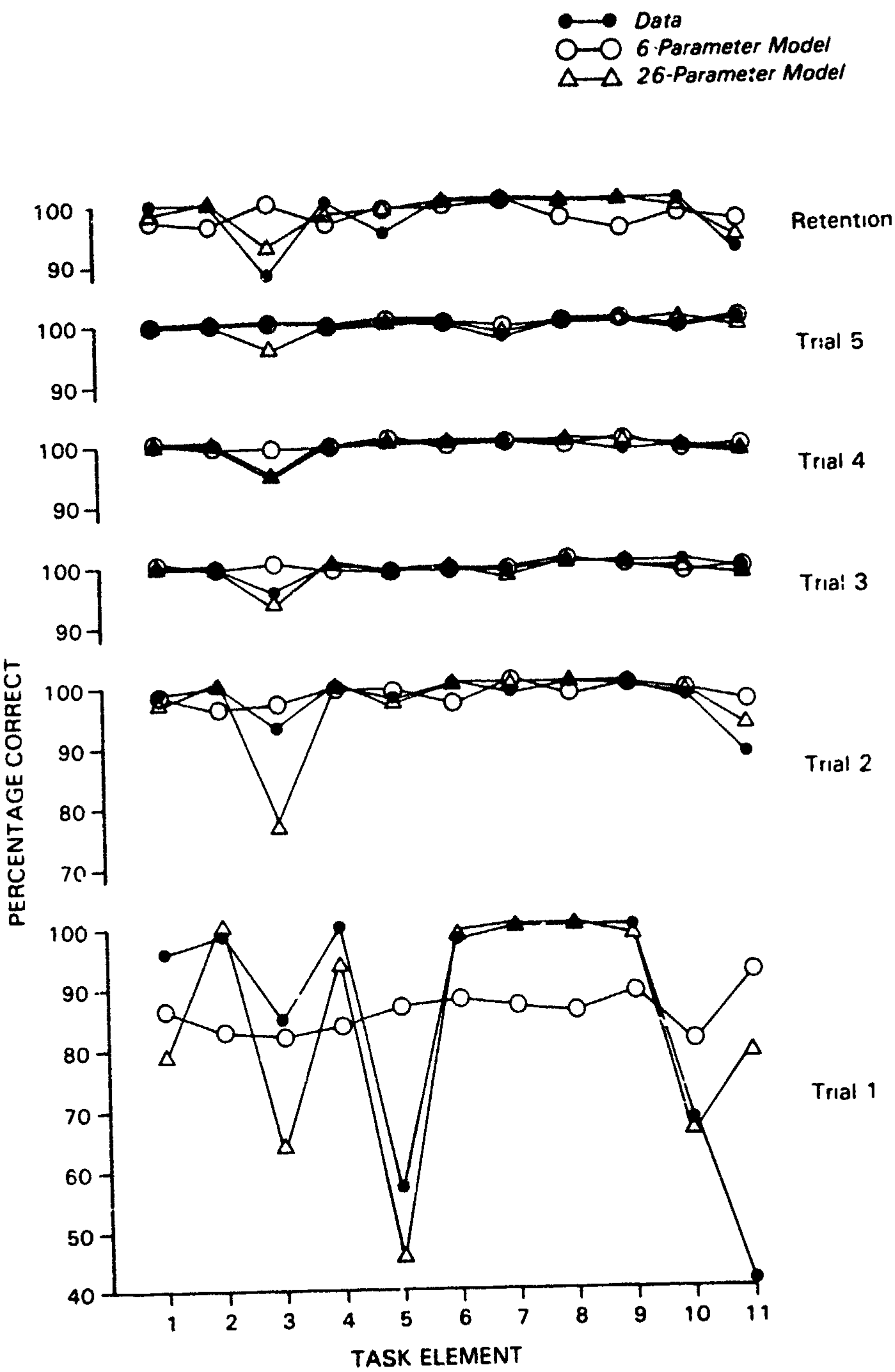

Figure 1

PREDICTED AND ACTUAL TASK-ELEMENT PERFORMANCE FOR LOADMG

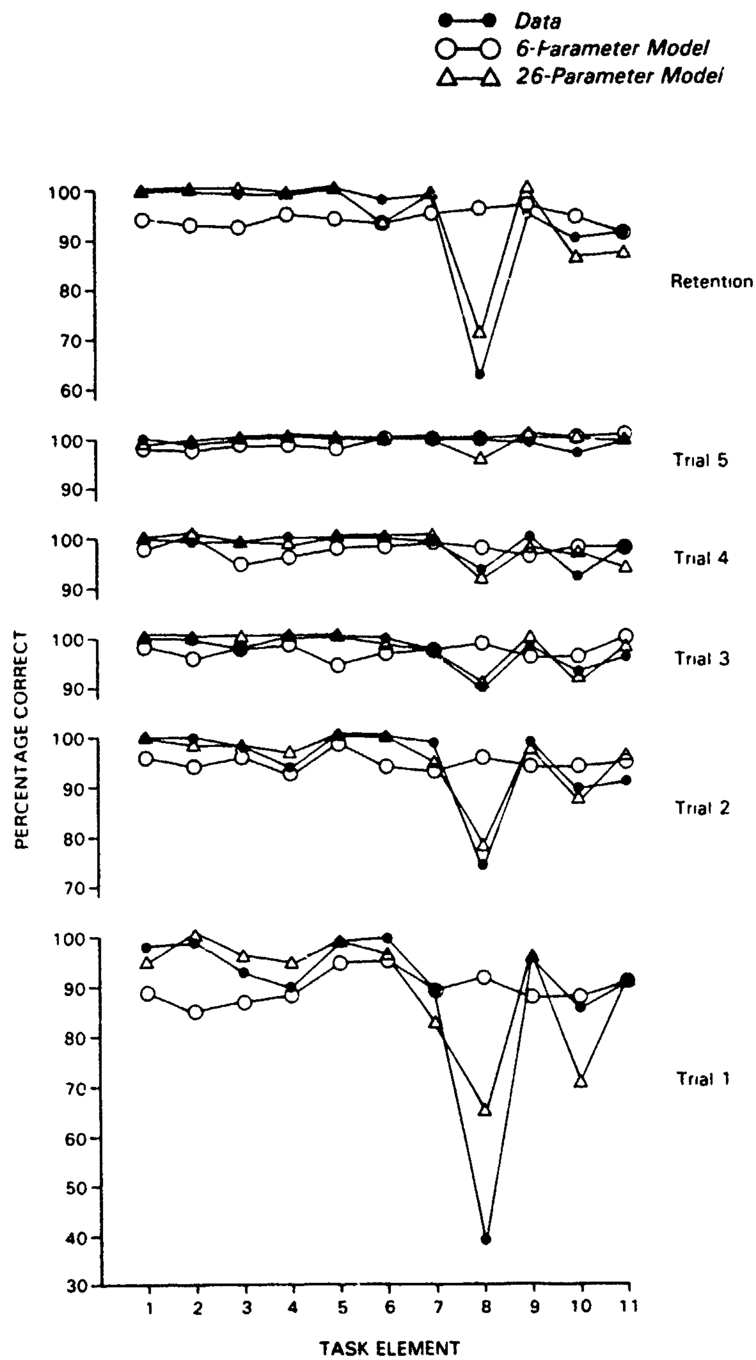

Figure 2

PREDICTED AND ACTUAL TASK-ELEMENT PERFORMANCE FOR STARTANK

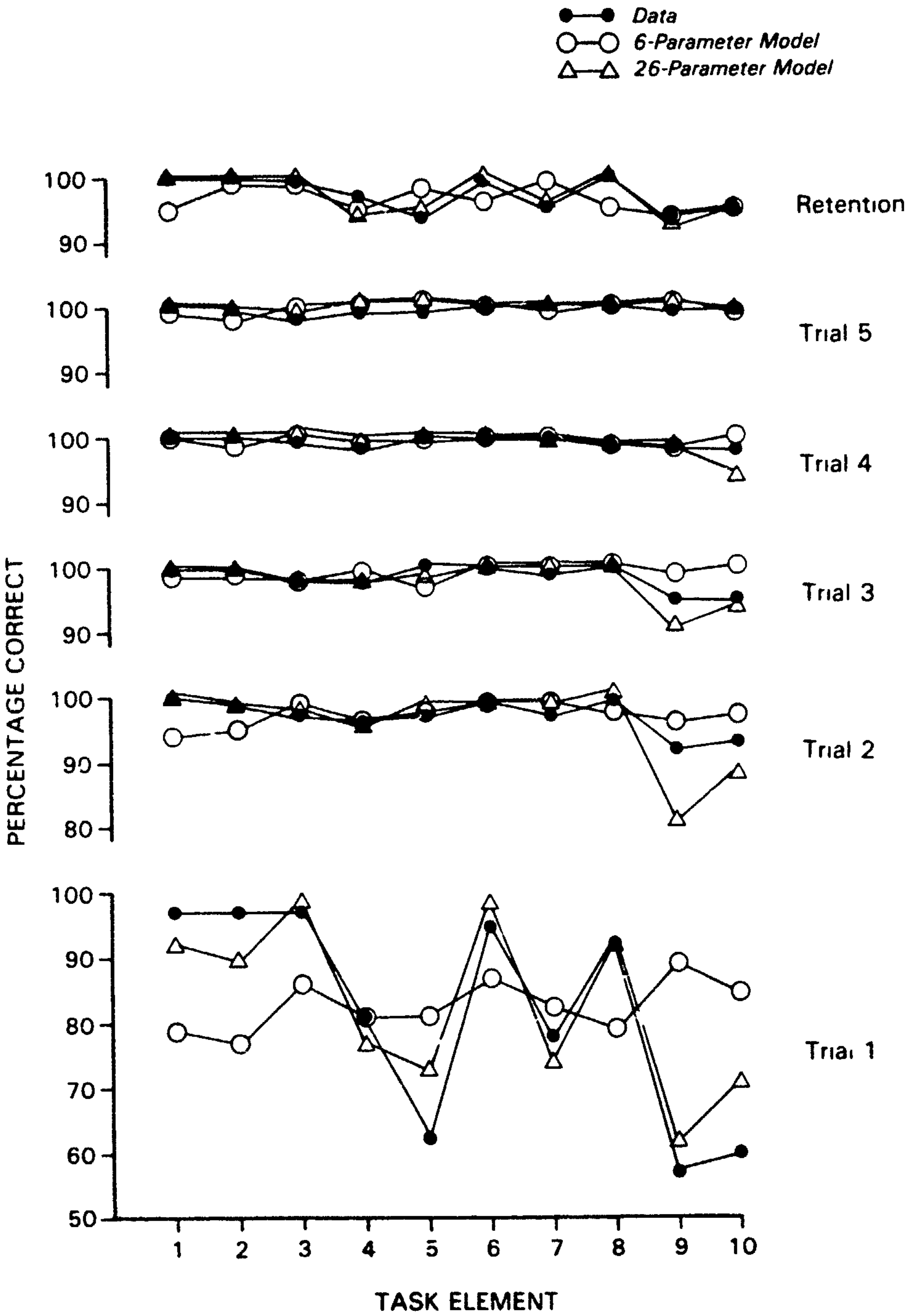

Figure 3

PREDICTED AND ACTUAL TASK-ELEMENT PERFORMANCE FOR STOPTANK

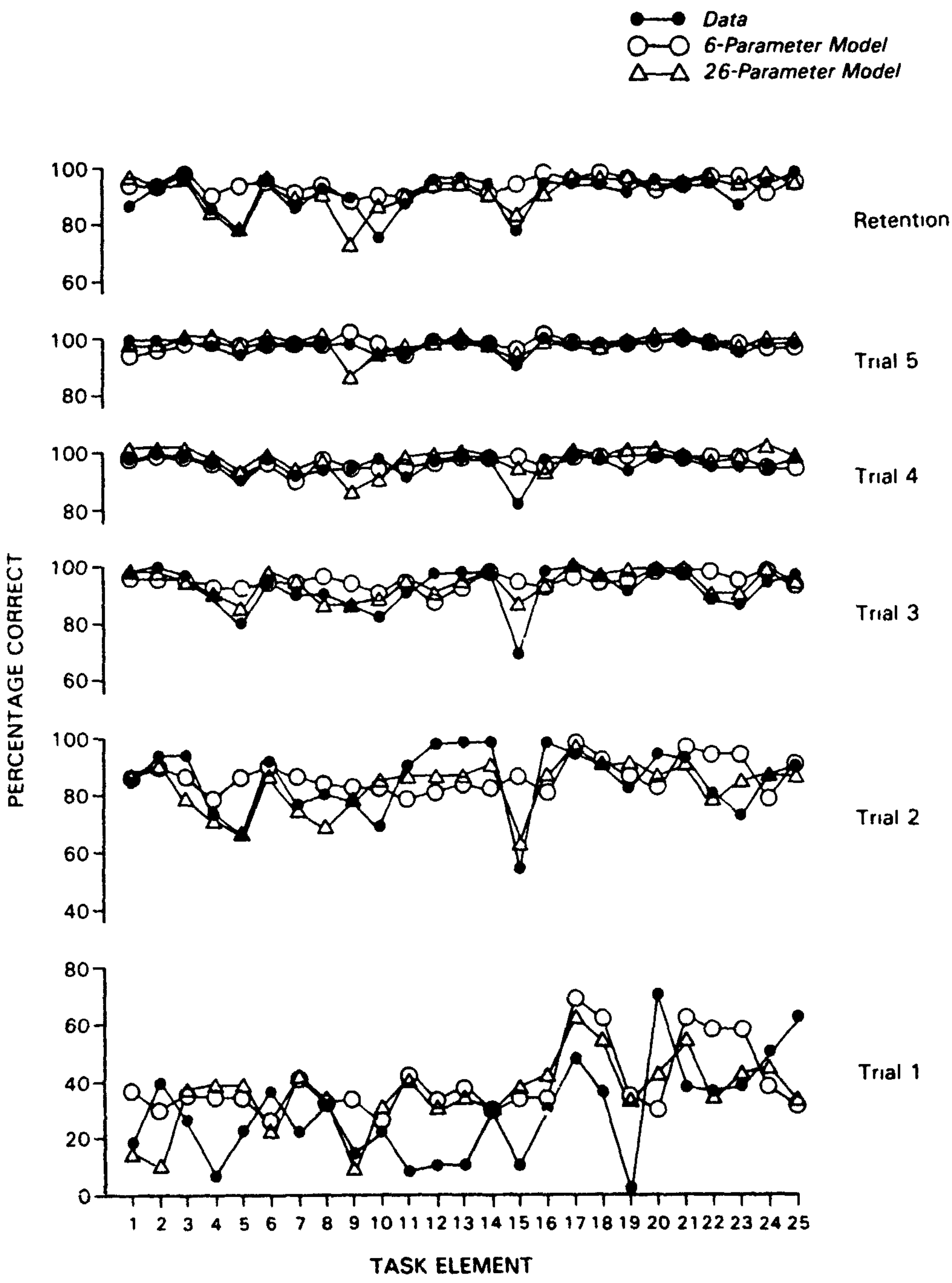

Figure 4
PREDICTED AND ACTUAL TASK-ELEMENT PERFORMANCE FOR GUNNERPF

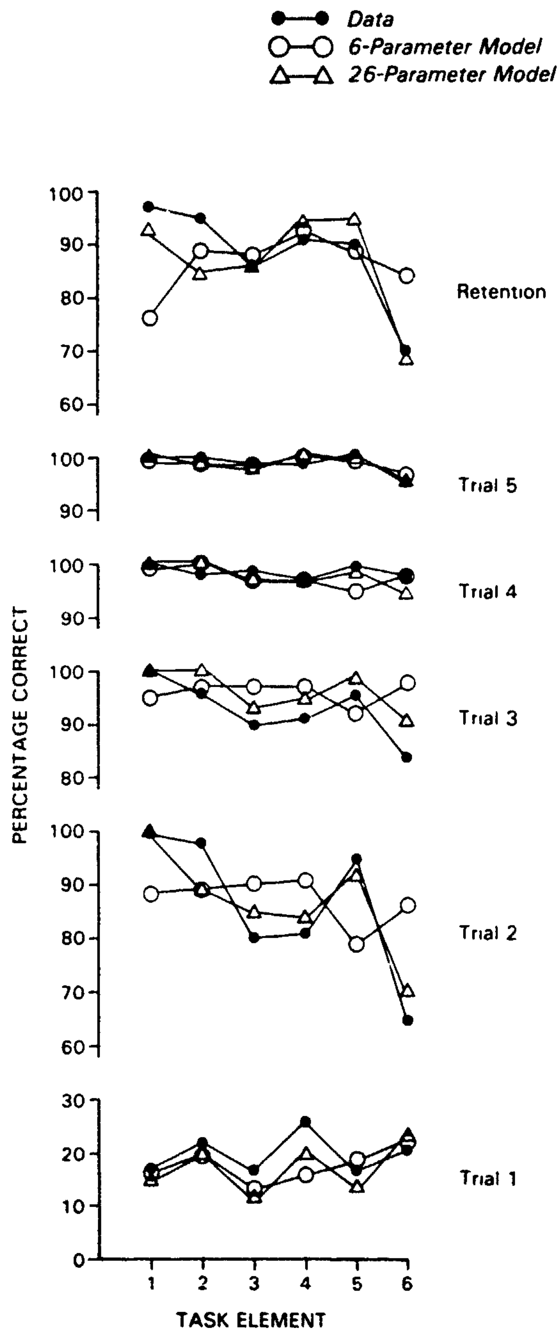

Figure 5
PREDICTED AND ACTUAL TASK-ELEMENT PERFORMANCE FOR LOADERPF

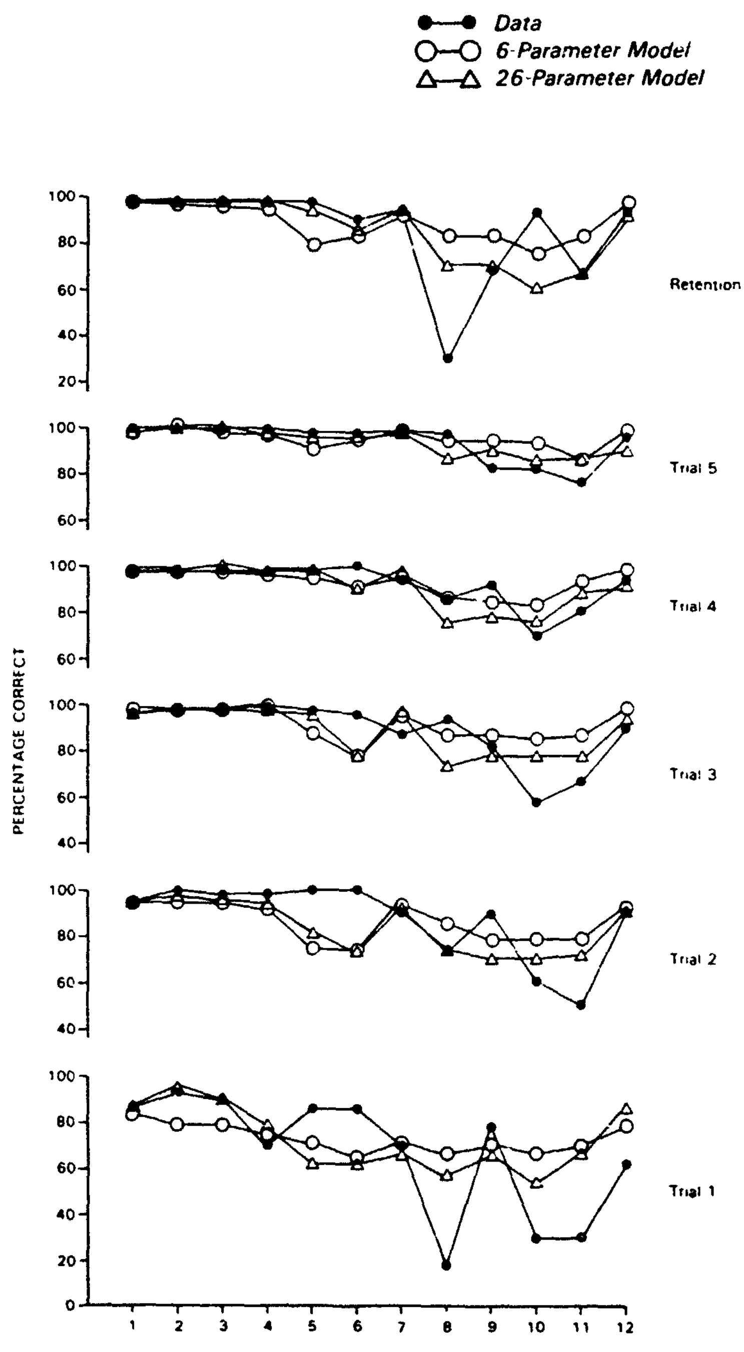

Figure 6

PREDICTED AND ACTUAL TASK-ELEMENT PERFORMANCE FOR PRECFIRE

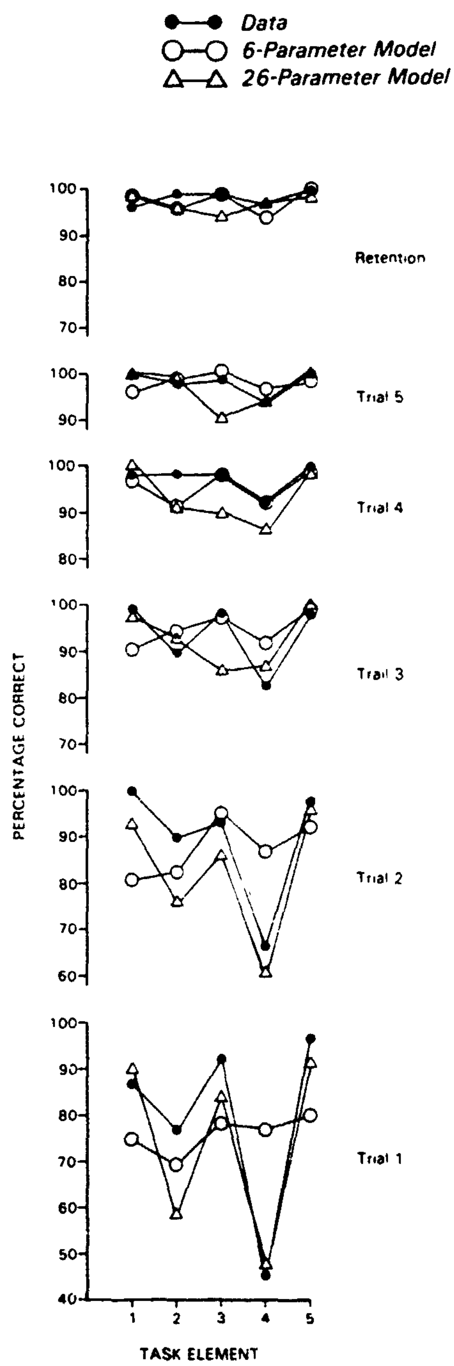

Figure 7

PREDICTED AND ACTUAL TASK-ELEMENT PERFORMANCE FOR RADIOMSG

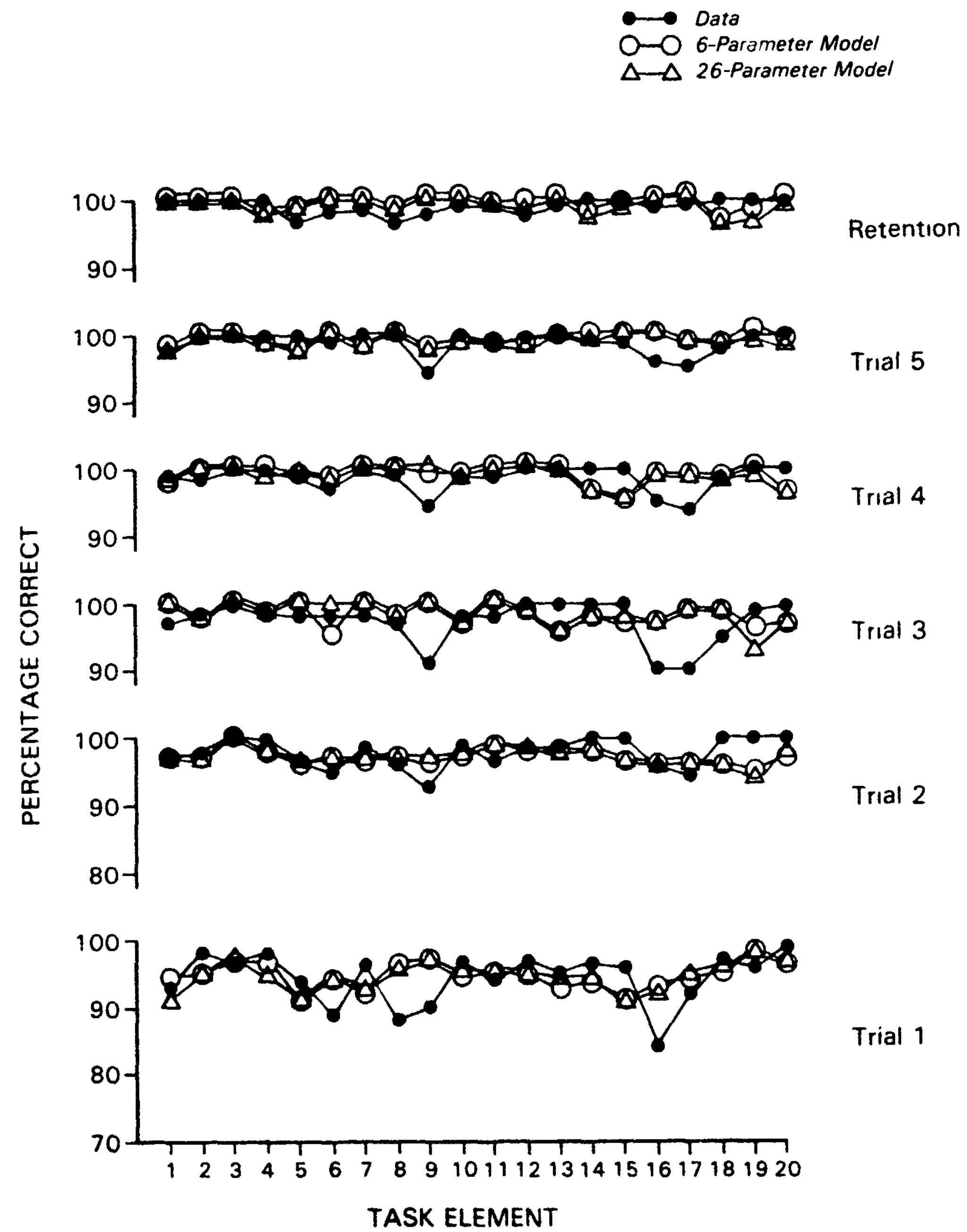

Figure 8

PREDICTED AND ACTUAL TASK-ELEMENT PERFORMANCE FOR SIGNALS

leading to different weights for tasks that are generally high on a characteristic, than for those tasks that are generally low on the characteristic. Third, even if task characteristics were measured on an interval scale, it would not be surprising if the relationship between the model parameters and task characteristics were not linear, or even monotonic. Certain task characteristics, such as the extent of aversive consequences, may have an intermediate value that produces greatest learning or retention. A linear approximation to this single-peaked function may be good for a single task, in which the range of values for the task characteristic is small, but different tasks would produce different linear relationships, which could differ greatly.

The ability of the task characteristics to predict differences between task elements from different tasks was tested using linear regression of the parameter values predicted by the task-characteristic model on the task ratings. This linear regression should be interpreted in light of the comments stated in the previous paragraph. Three of the four parameters, initial strength, strength asymptote, and retention proportion, cannot be compared across tasks, because strength is measured on an interval scale. Consequently, these parameters were transformed to three parameters which provide a basis for more meaningful comparisons. Initial strength and the strength asymptote were transformed by subtracting from them the value of the threshold for a correct response. If the strength required for a correct response is constant over task elements, then this transformed value may be meaningfully compared across tasks. The retention proportion was transformed by calculating the amount of decay, which is the difference between the predicted strength on the sixth trial and what would have been predicted if there were no decay during the retention interval. The learning rate was not transformed.

Independent variables for the analysis were the values on the five task characteristics and the number of steps in the task. The number of steps was included in the regression because it has been found to relate to learning and retention of procedural tasks (Shields et al., 1979; Knerr et al., 1982).

The results of the analysis (Table 4) indicate that the six independent variables account for a significant proportion of the variability of initial strength, learning rate, and strength asymptote. The number of steps in the task accounts for the most variance for each of these three model parameters; however, all skill ratings except whether the task is performed in combat are significantly related to at least one of the dependent variables. An increase in the number of steps in the tasks was associated with decreased initial strength and strength asymptote, and increased learning rate. Greater use of rules was associated with greater initial strength and retention, and

Table 4
Results of Regression Analysis of Task Characteristics Across Tasks

Dependent Variable	R^2	Independent Variable Weights: Number of Steps	Use of Rules	Aversive Consequences	Feedback	Interference	Performed In Combat
Initial Strength (Standardized)	0.40***	-0.040***	0.081*	0.015	0.053*	-0.045	-0.948
Learning Rate	0.35***	0.009***	0.007	-0.030***	-0.018***	-0.002	-0.079
Strenth Asymptote (Standardized)	0.24***	-0.078***	-0.181*	0.072	0.048	0.177***	0.427
Decay Amount	0.07	-0.014	-0.131*	0.042	-0.003	0.023	-0.421

*$p < .05$

**$p < .01$

***$p < .001$

lower strength asymptote. Aversive consequences were negatively related to learning rate among the eight tasks. Greater feedback led to greater initial strength, but lower learning rate. Finally, greater interference was associated with a greater strength asymptote.

Individual Differences

Individual differences were identified by Sticha et al. (1982) by applying a model developed for one set of soldiers to the data from another set of soldiers. The maximum-likelihood values for the parameters, estimated from the second set of soldiers, provided a significantly better account for those data than the parameters estimated from the first group for most of the tasks, indicating the existence of individual differences. Although they were significant, the size of these differences was relatively small.

In this analysis, the values of the four learning and retention model parameters are related to two measures of soldier aptitude, AFQT percentile and the Combat (CO) scale of the ASVAB. Each of the learning and retention parameters was assumed to be a linear combination of the two aptitude measures. The resulting model has 14 parameters: 3 thresholds; 3 constants for the linear equations for learning rate, strength asymptote, and retention proportion; and 8 weights representing the weights of 2 independent variables predicting 4 dependent variables. (Models for PRECFIRE and SIGNALS have one additional parameter, the constant for the equation describing the initial strength.) The model was limited to this size because of the time involved in parameter estimation, which is somewhat greater than the time required for models of task-element differences.

Table 5 shows goodness-of-fit for the individual difference models, and the improvement of these models over models which assume no individual differences (the basic six-paramenter model). The basic models are the same as shown in Table 2. However, the goodness-of-fit measures are not the same in the two tables, because soldiers for whom ASVAB scores were not available were eliminated from the analysis of individual differences. The results show significant improvement in predictions in four tasks; improvements correspond in magnitude to those reported by Sticha et al. (1982).

The weights were standardized by multiplying them by the standard deviation of AFQT (18.60) and CO (12.88) scores for all soldiers in the sample. The standardized weights (Table 6) indicate somewhat more consistency across tasks than was present in the weights for task-element characteristics. For example, CO has a positive weight on the initial strength for all of the tasks, indicating that those soldiers who are high in this aptitude learn more from the formal training that occurred before the training trials conducted in the course of this study.

Table 5
Goodness-of-Fit for Individual Difference Models

Task	Negative Log Likelihood: Single-Value Model	Negative Log Likelihood: Individual Difference Model	Chi-Square for Improvement[a]
LOADMG	1584.82	1581.33	3.49
STARTANK	1729.41	1722.93	6.48
STOPTANK	1626.12	1610.76	15.36
GUNNERPF	14463.96	14366.96	97.00**
LOADERPF	1718.92	1709.33	9.59
PRECFIRE	7295.22	7217.44	77.78**
RADIOMSG	3565.96	3522.74	43.22**
SIGNALS	3209.14	3186.94	22.20*

[a] df=8

*$p < 0.01$

**$p < 0.001$

Table 6
Standardized AFQT and CO Weights

TASK	DEPENDENT VARIABLE	INDEPENDENT VARIABLE WEIGHTS AFQT	CO
LOADMG	Initial Strength	-0.041	0.087
	Learning Rate	0.008	0.021
	Asymptote	0.023	0.013
	Retention	0.009	-0.009
STARTANK	Initial Strength	-0.038	0.070
	Learning Rate	-0.026	0.002
	Asymptote	0.714	0.165
	Retention	0.017	-0.030
STOPTANK	Initial Strength	-0.168	0.240
	Learning Rate	-0.022	0.141
	Asymptote	0.169	-0.186
	Retention	0.017	0.004
GUNNERPF	Initial Strength	-0.106	0.136
	Learning Rate	0.085	-0.091
	Asymptote	-0.267	0.245
	Retention	0.035	0.0
LOADERPF	Initial Strength	0.139	0.024
	Learning Rate	0.047	-0.052
	Asymptote	-0.155	0.252
	Retention	0.027	-0.029
PRECFIRE	Initial Strength	-0.117	0.137
	Learning Rate	0.071	-0.020
	Asymptote	-0.367	0.315
	Retention	0.003	-0.005
RADIOMSG	Initial Strength	0.042	0.039
	Learning Rate	0.012	0.019
	Asymptote	0.126	-0.010
	Retention	-0.005	-0.032
SIGNALS	Initial Strength	-0.187	0.033
	Learning Rate	0.043	0.077
	Asymptote	0.026	-0.020
	Retention	0.0	0.0

In addition, the absolute value of the weight for either of the aptitude measures was greater in the prediction of learning rate than it was for the retention proportion (in 13 of 16 instances). This result would indicate that aptitude is more closely related to learning than retention, a finding that is consistent with previous research. The standard deviation of the dependent variables is unknown, and hence, rigorous comparison of weights across dependent variables is not possible. However, since both learning rate and retention proportion have values between 0 and 1, the standard deviations should be roughly comparable. The fact that retention proportions tend to be more extreme than learning rates may indicate that they have a lower standard deviation, and hence, partially explain the difference in weights.

Another striking pattern in Table 6 is that in 21 of 32 cases, the signs for weights for AFQT and CO have the opposite sign. One interpretation of this result is that when both composites are used, one acts as a suppressor. If this interpretation is correct, future research should either select one of these two composites, or combine them to form a single predictor which may be more reliable.

Discussion

The results of this research bear on the validity of the model of procedural learning and performance, the application of analytic models of issues regarding the acquisition and retention of skills, and the specific issues addressed in this analysis, task-element and individual differences in learning and retention.

Model Validity

The results of Sticha et al. (1982) showed the capability of the analytic models embodied in the SAINT framework to describe general characteristics of procedural learning and retention. The model provided the ability to predict performance accuracy and speed at the whole-task level. In addition, the models predicted the average accuracy at the task-element level. The models developed in the present analysis extend the results of Sticha et al. to predict differences between task elements or individuals. Thus, the current models provide considerably greater detail than the original models.

However, it should be realized that that models developed in this analysis are largely exploratory. Sticha et al. (1982) validated their models by applying the parameters estimated from one set of subjects to the data from another set of subjects. Since the models developed in this analysis contained considerably more parameters than the simpler models which do not consider task-element or individual differences, the data were not divided into model development and cross validation groups, so that more stable parameter estimates could be obtained. Thus, the results of this research should be interpreted with the same care that is required for all "correlational" analyses. In addition, the results of this analysis need to be confirmed with replication studies, or analyses of other acquisition and retention data.

Methodological Issues

A major purpose of this report is to illustrate the application of mathematical models to investigate issues regarding acquisition and retention of complex, military skills. The application to task-element and individual differences has illustrated some of the aspects that characterize the methodology and distinguish it from more traditional methods.

The most obvious advantage of the model-based analysis is that it gives the researcher the ability to distinguish several components of learning and retention, such as initial level of learning, learning rate, and limits to learning. This increased level of analytical detail allows the researcher to localize

the effect of experimental variables to specific theoretical constructs. On the other hand, the increased theoretical complexity makes it more difficult to derive simple, general conclusions from experimental results. Whereas a researcher may attempt to make a direct generalization from the results of an empirical study, a model-based analysis will not allow such simple extrapolation. However, a model may be used to predict performance in any specific situation if it represents enough details about the situation.

The chief problem with the analytical methods described in this report is the time and resources that their application requires. The actual time (and cost) required for parameter estimation depends on the specific computer on which the analysis is being conducted, the optimization routine being used, the complexity of the model being tested, and the efficiency of the routine used to calculate goodness-of-fit. For this reason, it is difficult to estimate the cost or time required for parameter estimation for a particular application. On the other hand, it is clear that the methods described in this section involve a considerably greater committment of analytical resources than alternatives such as regression or analysis of variance. The time required to find optimal parameter values makes it difficult, if not infeasible, to do analyses, analogous to stepwise regression, that require repeated application of the optimization procedures.

There are a number of ways in which this problem may be addressed in future analyses. Of course, the simplest way would be to use more efficient estimation procedures. In trying different approaches to some of the problems addressed in this report and by Sticha et al. (1982), order of magnitude differences were often obtained in speed estimation of parameters between different procedures. If the most efficient methods for problems of this type could be determined, the time saved may be sufficient to allow more complex analyses.

An alternative to the analysis presented in this report would be to estimate the learning and retention parameters separately for each task element, and then use regression analysis to model the differences in parameter values between task elements. A disadvantage of this approach is that it requires a considerable amount of data to provide accurate parameter estimates at the task-element level; the amount of data in the current research would probably be near the lower limit for which the method could be applied. The analogous method for investigating individual differences by estimating parameters separately for each individual would probably be infeasible because of the difficulty of getting a large enough number of tasks to estimate individual learning parameters. The major advantage of this alternative is that it allows the powerful and simple methods of linear regression to be applied for

exploratory analysis rather than the much slower, iterative optimization routine.

Task-Element Differences

Results of the interference ratings appear to indicate that high interference levels are associated with higher rates of initial learning for some tasks. One possible reason for this result is that the OSUT instructors know which parts of the tasks will produce performance problems (e.g., those that are performed differently in similar tasks or in an operational unit) and, therefore, emphasize them during formal training. However, certain caveats apply; these pertain to the rating instruments and the potential fitting of error.

The ratings for interference, and the other task-element characteristics, were developed for this research and do not have the benefit of reliability and validity research. Improved measures might show results that are more consistent across tasks and with theoretical formulations. Some ways to improve the measurement of task-element characteristics are naturalistic observation of task performance and video taping of the performance. If ratings continue to be used, they can be refined by using scaling techniques, such as forced distributions or behaviorally anchored scales.

The number of task elements in some of the tasks was close to the number of task characteristics, and some fitting of error variance may result. The fact that the fit of the models is not as good for longer tasks (Figures 1-8) suggests that fitting of error is occurring. Tasks with especially high or low numbers of elements did not show consistent results regarding task characteristics. Tasks with interference index weights in the direction opposite to that expected were tasks in the middle range of length. These effects remain to be tested with improved data collection for the task characteristics.

Task Differences

The effect of the number of steps and the skill ratings on the learning parameters may be interpreted in light of the nature of the learning model. According to this model, learning on any trial is proportional to the amount to be learned and the learning rate. The amount to be learned is the difference between the current strength and the strength asymptote. Examination of the results shown in Table 4 shows that for all but one of the independent variables (amount of feedback), variables which increase the amount to be learned decrease the learning rate. This pattern of results suggests that the increase in strength brought about by a single training trial is a single-peaked function of the skill ratings; that is, there is a value of the ratings which maximizes the strength increase, depending on the

current strength. This result will be illustrated for the case of the number of steps.

The weights shown in Table 4 indicate that increasing the length of a task by one step, should decrease the amount to be learned as the soldier comes into the experiment (by 0.038) by decreasing both the initial strength (by 0.040) and the strength asymptote (by 0.078). In addition, the increase will lead to an increase in the learning rate (by 0.009). If the amount to be learned is high and the learning rate is low, the overall effect of adding a step to the task will be to increase the degree of learning that occurs on a single (or fixed number) of trials. Addition of a second task element should have a smaller effect, because the amount to learn has been lowered and learning rate increased. As more steps are added, the learning rate will become sufficiently high, so that making the task any longer will decrease the effectiveness of a single trial. Thus, for a fixed number of trials, there should be a task length which produces optimal learning.

The existence of single-peaked relationships between task characteristics, and the effectiveness of a fixed number of training trials may help to explain why different researchers may find different relationships between task characteristics and degree of learning. In addition, the results can help us understand why learning experiments may produce different results depending on the number of trials. For a small number of trials, the learning rate is smaller, and hence, increases in the number of steps (or other task characteristic) will increase learning. For a larger number of trials, the learning rate is larger, and hence, increases in the task characteristic will lead to decreased learning. In this case, the modeling results have the potential of explaining seemingly contradictory empirical findings.

Individual Differences

The aptitude measures considered did not improve model prediction to the extent of the task-element differences. This result may be caused, in part, by the fact that there are more subjects than there are task elements, and hence, the prediction of task-element differences is an easier task. Consistent with this interpretation, the weights of the aptitude measures were more consistent across task than were those of the task-element characteristics. Care should be taken in interpreting the results of the individual difference models, because of the possibility that one of the aptitude measures is acting as a suppressor.

Summary and Conclusions

The research described both in this report and elsewhere (Knerr et al., 1982; Sticha, 1982; Sticha et al., 1982) is focused on the development, validation, and application of mathematical models of procedural learning and retention. Both the progress that was made and the work that remains are substantial in these three activities.

Model development. The major accomplishments of this research are the development of integrated models of procedural learning and retention, and the incorporation of these models within a complex performance simulation model. The model was shown to predict accurately improvements in overall performance that occur during training and decay in performance that happens shortly after training is completed (Sticha et al., 1982). In addition, the models may be extended to predict learning and retention differences among task elements or individuals.

The process of estimating the values of model parameters still requires that considerable effort be applied to data collection and analysis. Development of an approach that allows the training researcher or training manager to estimate parameter values without extensive data collection is critical to the eventual success of the modeling approach. This report has described an approach based on ratings of task elements on several characteristics. The results are encouraging; however, further theoretical insights, methodological advancements, and data analysis are required to develop and validate methods for predicting model parameters.

Model validation. Because it was possible to separate the psychological models from the performance simulation for the purposes of model validation, it was possible to conduct a much more rigorous and complete model validation than is typical for simulation models of similar scope and complexity. In particular, it was possible to determine optimal values for model parameters, and test hypotheses about model adequacy using formal statistical procedures. This approach to model validation has considerable advantages over less rigorous approaches based on sensitivity analyses. Thus, design of future validation research should consider the substantial validation effort that has already taken place.

However, there is a need for further empirical validation of the retention component of the model. The two samples of data from OSUT and from the operational unit seem to be giving different pictures of the changes in the strength of the memory trace that occur after initial training. On the one hand, considerable forgetting occurs in the one-month retention interval for the OSUT soldiers. On the other hand, performance is constant over the interval from three months to two years

investigated in the operational unit. Although these results are consistent with the retention functions considered in the models, information from the first three months after training, which is critical to assessing the shape of the retention function, is unavailable. Since soldiers who have graduated from OSUT within three months are generally unavailable for study, data will be difficult to obtain. One approach to obtaining retention data involves use of a within-subjects design. In this design soldiers in an operational unit who differ in the time since OSUT would be trained on a task and tested after a retention interval. The loss in performance during the retention interval would allow for the estimation of decay parameters and validation of the retention model.

Model application. This report was intended to illustrate how mathematical models could be applied to investigate issues regarding the acquisition and retention of military skills. The use of mathematical models for data analysis represents the most immediate application of models. Other applications involve the development of a system to support the needs of training researchers and training managers. Such a system would (1) organize the results of learning and retention experiments for researchers and guide in the design and interpretation of new research, and (2) make predictions for managers regarding the effects of various schedules of initial and refresher training on performance levels. Although such an application requires further model development and validation, in addition to system design and development, it is critical to assess the needs of managers and researchers for such a system early, so that the maximum benefit may be gained from future research.

References

Annett, J. Skill loss: A review of the literature and recommendation for research. Warwick, England: University of Warwick, September 1977.

Dressell, J. D. Mnemonically enhanced training (Working Paper). Alexandria, VA: U. S. Army Research Institute, February 1980.

Goldberg, S. L., Drillings, M., & Dressel, J. D. Mastery training: Effect on skill retention (Technical Report 513). Alexandria, VA: U. S. Army Research Institute, March 1981.

Hagman, J. D. Effects of training task repetition on retention and transfer of maintenance Skill (Research Report). Alexandria, VA: U. S. Army Research Institute, May 1980(a).

Hagman, J. D. Effects of training schedule and equipment variety on retention and transfer of maintenance skill (Research Report). Alexandria, VA: Army Research Institute, November 1980(b).

Hull, C. L. Principles of behavior. New York: Appleton-Century-Crofts, 1943.

Hull, C. L. A behavior system: An introduction to behavior theory concerning the individual organism. New Haven, CN: Yale University Press, 1952.

Knerr, C. M., Berger, D. C., & Popelka, B. A. Sustaining team performance: A systems model (Final Report). Springfield, VA: Litton Mellonics, 1980.

Knerr, C. M., Harris, J. H., O'Brien, B. K., Sticha, P. J., & Goldberg, S. L. Armor procedural skills learning and retention (DDI/PR 82-13-334). McLean, VA: Decisions and Designs, Inc., October 1982.

McCluskey, M. R., Hiller, J. H., Bloom, R. D., & Whitmarsh, P. J. Skill decay of sixteen common tasks for MOS 11B and 11C (Final Report). Alexandria, VA: Army Research Institute, November 1978. (Draft)

Osborn, W. C., Campbell, C. H., & Harris, J. H. The retention of tank crewman skills (Research Report 1234). Alexandria, VA: U. S. Army Research Institute, December 1979.

Rose, A. M., McLaughlin, D. H., Felker, D. B., & Hagman, J. D. Retention of soldiering skills review of recent ARI research (Technical Report 530). Alexandria, VA: Army Research Institute, in press.

Schendel, J. D., Shields, J. L., & Katz, M. S. Retention of motor skills: A review (Technical Paper 313). Alexandria, VA: U. S. Army Research Institute, September 1978.

Shields, J. L., Goldberg, S. L., & Dressel, J. D. Retention of basic soldiering skills (Research Report 1225). Alexandria, VA: U. S. Army Research Institute, September 1979.

Sticha, P. J. Review of analytical models of procedural learning and performance (DDI/PR 82-14-334). McLean, VA: Decisions and Designs, Inc., November 1982.

Sticha, P. J., Edwards, T. D., & Patterson, J. F. An analytical model of learning and performance of armor procedures (DDI/PR 82-15-334). McLean, VA: Decisions and Designs, Inc., November 1982.

Sullivan, D. J., Casey, R. J., & Hebin, J.M. Acquisition and retention of cognitive versus perceptually oriented training materials (Technical Report). Alexandria, VA: Army Research Institute, October 1978.

U. S. Army Field Artillery School. Weapon system training effectiveness analysis: The forward observer (Phase IA Baseline, Vol. 1). Fort Sill, OK: Author, 1977.

U. S. Army TRADOC Systems Analysis Activity (TRASANA). REDEYE weapons system training effectiveness analysis (TRASANA Technical Report No. 21-77). White Sands, NM: Author, 1977.

Vineberg, R. A study of the retention of skills and knowledge acquired in basic training (HumRRO Technical Report 75-10). Alexandria, VA: Human Resources Research Organization, June 1975.

Wheaton, G. R., Rose, A. M., Fingerman, P. W., Korotkin, A. L.; Holding, R. F. & Mirabella, A. Evaluation of the effectiveness of training devices: Literature review and preliminary model (Research Memorandum 76-6). Alexandria, VA: U. S. Army Research Institute, April 1976.

Wickelgren, W. A. Single-trace fragility theory of memory dynamics. Memory and Cognition, 1974(a), 2, 775-780.

Wickelgren, W. A. Strength/resistance theory of the dynamics of memory storage. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Sugges (Eds.), Contemporary developments in mathematical psychology (Vol. 1). San Francisco: W.H. Freeman, 1974(b).

Wortman, D. B., Duket, S. D., Seifert, D. J., Hann, R. L., & Chubb, G. P. Simulation using SAINT: A user-oriented instruction manual (AMRL-TR-77-61). Dayton, OH: Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, 1978(a).

Wortman, D. B., Duket, S. D., Seifert, D. J., Hann, R. L., & Chubb, G. P. The SAINT users manual (AMRL-TR-77-62). Dayton, OH: Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, 1978(b).